

楽譜データに基づく楽曲の特徴量抽出に関する研究

著者	米倉 弥伸
出版者	法政大学大学院理工学研究科
雑誌名	法政大学大学院紀要．理工学・工学研究科編
巻	61
ページ	1-5
発行年	2020-03-24
URL	http://doi.org/10.15002/00022910

楽譜データに基づく楽曲の特徴量抽出に関する研究

RESEARCH OF EXTRACTING THE CHARACTERISTICS OF MUSIC BASED ON SCORE DATA

米倉 弥伸

Hisanobu Yonekura

指導教員 藤井 章博

法政大学大学院理工学研究科応用情報工学専攻修士課程

There are various approaches to music analysis. The reason for conducting this research is to consider the feasibility of a system that contributes to the support of human creators' music composition activities by analyzing the music, accumulating the results of the analysis as a knowledge graph, and developing it. In this research, we perform music analysis using MusicXML structured data and accumulating the results of the analysis.

Key Words : MusicXML, music analysis

1. はじめに

一般的に、音楽などの芸術作品での分析においては、その分野での専門知識をもとにして主観的な評価をされることが多く、分析する人によって様々な評価がされる。また、楽曲の特徴としては、例えばバロック時代の舞踏に供する楽曲であれば、メヌエットなどの様式に即した楽曲の構成がなされていて、作曲者の創作意図の前提となっている。創作意図を汲み取ることは作曲者や演奏者にとって楽曲を理解するためには重要な情報となる。そのため、音楽の分野において主観的な評価ではなく客観的な評価をできるようにすることは有意義なことであると考えられる。

これまで、楽曲の分析の取り組みとしては多種多様なものがある。[1]そのような状況の中で本研究を行う理由としては、楽曲を分析し、分析の成果を知識グラフという構造化したデータとして蓄積し、これを発展させることで人間の創作者による作曲活動の支援に資するシステムの実現の可能性を検討したいという点である。すなわち、音楽学の分野で「アナリゼ」と呼ばれる行為を支援し、演奏者や作曲者が創造性を発揮するための支援を行う方法を検討したいということである。

本研究では、楽曲の構造に着目しての分析を行い、分析した結果を知識グラフとして蓄積することを目的とする。MusicXML 形式の構造化データを利用し、複数の共通した様式から構成される組曲についての分析を行う。具体的には、各楽曲の旋律、調、音名、音価の関係を分析し、1つの組曲を通じて現れる特徴や、曲種ごとの類似性等の特徴量を抽出することができるかについて調査する。これにより数理的構造のみから楽曲をどの程度示すことが

できるかについて検討する。

2. 分析対象楽曲

本研究では、分析対象楽曲としてバッハが作成した無伴奏チェロ組曲第1番～第6番 (BMW1007～BMW1012) を編曲した、「ギターのための無伴奏チェロ組曲全曲集」[2]に収録されている36曲を扱った。表1に分析対象楽曲の一覧を示す。

表1 分析対象楽曲一覧

第1番 (ト長調)	第2番 (ニ短調)	第3番 (ハ長調)
前奏曲 アルマンド クーラント サラバンド メヌエット ジグ	前奏曲 アルマンド クーラント サラバンド メヌエット ジグ	前奏曲 アルマンド クーラント サラバンド ブレー ジグ
第4番 (変ホ長調)	第5番 (ハ短調)	第6番 (ニ長調)
前奏曲 アルマンド クーラント サラバンド ブレー ジグ	前奏曲 アルマンド クーラント サラバンド ガヴォット ジグ	前奏曲 アルマンド クーラント サラバンド ガヴォット ジグ

分析対象とする組曲は、古典的な舞踏曲を数種集めたものに前奏曲を加えたものであり、1曲目に前奏曲、2曲目

にアルマンド, 3 曲目にクーラント, 4 曲目にサラバンド, 5 曲目にメヌエット, あるいはほかの舞曲曲, 6 曲目にジグとなっている。1 つの組曲はそれぞれ 6 曲で構成され, 組曲 6 つで合計 36 曲となる。解説書[3]による舞曲曲それぞれの特徴を表 2 に示す。

表 2 各楽曲の特徴

前奏曲	即興的な自由な形式なものが多い
アルマンド	4 分の 4 拍子で中庸な速度をもった舞曲曲
クーラント	4 分の 2 拍子で力強い舞曲曲
サラバンド	4 分の 3 拍子, あるいは 2 分の 3 拍子による極めて緩やかな舞曲曲
メヌエット	4 分の 3 拍子で典雅な舞曲曲
ジグ	8 分の 3 拍子, あるいは 8 分の 6 拍子, または 8 分の 12 拍子などの極めて迅速で活発な舞曲曲

3. MusicXML

MusicXML は, 2004 年にアメリカの Recordare 社が制定した XML ベースの楽譜表記のためのフォーマットである。[4]MusicXML が制定される以前では, 音楽ソフトそれぞれが独自のフォーマットで開発されていたが, MusicXML という統一規格によりソフト間の垣根を超えて楽譜データのやり取りを行うことが可能となった。MusicXML 形式の記述例を図 1 に示す。

```
<measure number="2" width="615">
  <print/>
  <note>
    <pitch>
      <step>A</step>
      <octave>3</octave>
    </pitch>
    <duration>4</duration>
    <voice>2</voice>
    <type>quarter</type>
    <stem>down</stem>
    <notations/>
  </note>
```

図 1 MusicXML での記述例

図 1 に示したように, note タグは子として pitch タグや duration タグなどを持ち, pitch タグは子として step タグや octave タグを持つなど MusicXML は木構造を持つ要素の集合として表記される。また, step タグは音名, octave タグはオクターブ, duration タグは音価などそれぞれのタグが楽譜の構成要素を示す。

4. 分析方法

(1) 主旋律の準備

はじめに, 楽譜作成ソフト「KAWAI スコアメーカー 9pro」を用いて紙媒体の楽譜をスキャナで読み込み,

MusicXML 形式でエクスポートする。この時, 読み込んだ後の楽譜に認識ミスがあればエクスポートする前に修正を行う。その後, python のプログラムを用いて MusicXML から主旋律データを取り出す。取り出した主旋律データの例を図 2 に示す。

```
[[('a', 0.25), ('e', 0.25), ('cc#', 0.25), ('b', 0.25), ('cc#', 0.25), ('e', 0.25), ('cc#', 0.25), ('e', 0.25), ('a', 0.25), ('e', 0.25), ('cc#', 0.25), ('b', 0.25), ('cc#', 0.25), ('e', 0.25), ('cc#', 0.25), ('e', 0.25)],
[('a', 0.25), ('f#', 0.25), ('dd', 0.25), ('cc#', 0.25), ('dd', 0.25), ('f#', 0.25), ('dd', 0.25), ('f#', 0.25), ('a', 0.25), ('f#', 0.25), ('dd', 0.25), ('cc#', 0.25), ('dd', 0.25), ('f#', 0.25), ('dd', 0.25), ('f#', 0.25)],
[('a', 0.25), ('g#', 0.25), ('dd', 0.25), ('cc#', 0.25), ('dd', 0.25), ('g#', 0.25), ('dd', 0.25), ('g#', 0.25), ('a', 0.25), ('g#', 0.25), ('dd', 0.25), ('cc#', 0.25), ('dd', 0.25), ('g#', 0.25), ('dd', 0.25), ('g#', 0.25)],
[('dd', 0.25), ('g#', 0.25)]]
```

図 2 主旋律データの例(組曲第 1 番前奏曲冒頭 3 小節)

図 2 において, 主旋律データは音名と音価で 1 つのタプルとし, 1 小節ごとに 1 つのリストとしている。また, オクターブを示すために中央の「C」の音を基準として基準より 1 オクターブ高い音は音名を 2 つ重ね, 1 オクターブ低い音が小文字になるようにしている。

(2) 構成音の分析

構成音の分析として, 音階と音価の分析を行う。楽曲にはハ長調やヘ長調などのように調があり, 調ごとに音階がある。[5]図 3 に音階の例を示す。



図 3 音階の例

音階は, 最初の音から第 I 音, 第 II 音, 第 III 音, と順に呼ぶ。音階それぞれの音の特徴を表 3 に示す。

表 3 音階の各音の特徴

第 I 音	主音。音階の基点となる最も重要な音。
第 II 音	上主音。主音の 1 つ上の音。
第 III 音	中音。主音と属音の中間の音。
第 IV 音	下属音。主音と属音の補助役となる音。
第 V 音	属音。主音の次に重要な音。
第 VI 音	下中音。主音から下方の下属音との中間の音。
第 VII 音	導音。主音を導く音。

また, 音価とは, ある音符や休符の楽譜上の長さである。本研究では 1 拍分の音価を 1 とする。

音階の分析として, 主旋律データの音名から音階それぞれの音をカウントし, 第何音がどのくらいの割合を占めているかを分析する。また, 音価の分析として, 主旋律データからそれぞれの音価の音をカウントし, どの音価の音がどのくらいの割合を占めているかを分析する。

(3) n-gram を用いた構成音の分析

n-gram とは、任意の文字列や文書を連続した n 個の文字や単語で分割する方法である。例として、「AABDCA」という音階の文字列の時、次のように分割できる。

bi-gram (n=2) : 「AA」「AB」「BD」「DC」「CA」

tri-gram (n=3) : 「AAB」「ABD」「BDC」「DCA」

音階、音価の文字列に対し n-gram を用いてクラスタリングを利用した分析を行う。

(4) フレーズの分析

本研究では、連続する音符を区切ってメロディーの1つの単位を「フレーズ」と定義して分析対象とする。フレーズを厳密に定義することは一般的には困難であるが、ここでは舞曲曲であるという特徴を捉えて4小節を1フレーズとする。それぞれのフレーズは「テーマ x (A, B, …)」「テーマの展開 x'」のいずれかに属すると定める。判別には、音階からなる文字列の Levenshtein 距離を利用する。

Levenshtein 距離とは2つの文字列がどの程度異なっているかを示す距離の1種である。具体的には、文字単位で挿入・削除・置換の操作を行い、片方の文字列をもう一方の文字列に変形するために必要最小回数となる。例として、文字列Aが「ABCDEF」、文字列Bが「BBDEEF」の時、Levenshtein 距離は以下になる。

「B」を「A」に置換：「ABDEEF」

「C」を挿入：「ABCDEEF」

「E」を削除：「ABCDEF」

→Levenshtein 距離：3

2つのフレーズ a と b があった時、a と b の Levenshtein 距離を editdist (a, b) とすると a と b の類似度は次の式で求める。

$$sim(a, b) = 1 - \frac{editdist(a, b)}{Max(|a|, |b|)} \quad (1)$$

あるフレーズに対し、テーマとの類似度が0.4以上または1つ前のフレーズとの類似度が0.4以上の時をテーマの展開、テーマとの類似度が0.4より下の時は新たなテーマとしてフレーズの分析を行う。

5. 分析結果

(1) 構成音の分析結果

構成音の分析結果の例として、組曲第1番の音階の分析結果を図4に、各組曲の前奏曲の音価の分析結果を図5に、ジグの音価についての分析結果を図6に示す。

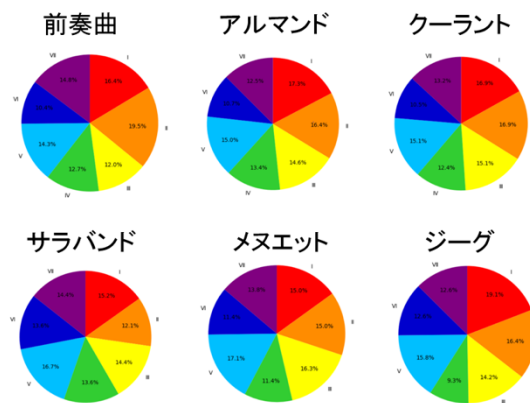


図4 組曲第1番の音階の分析結果

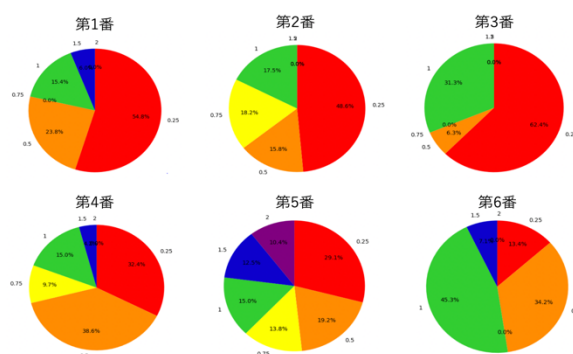


図5 各組曲の前奏曲の音価の分析結果

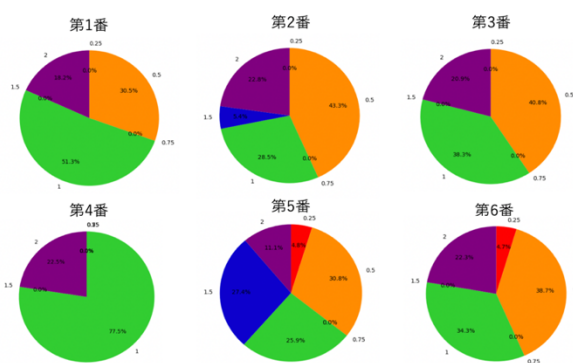


図6 各組曲のジグの音価の分析結果

音階については、組曲第1番のジグや組曲第6番のサラバンドとガヴオットを除き、割合が最も高い音と低い音で2倍の差がつくような大きな偏りは見られなかった。このことから、主音や属音のように重要な音ほど割合が高いということはなく、組曲全体を通じてそれぞれの音が満遍なく使われていると考えられる。

音価については、図5に示したように前奏曲では音価の小さい音が多く使われているものもあれば音価の大きい音が多く使われているものもあるという結果となった。また、サラバンドではほかの楽曲と比べて音価の大きい音の割合が高い結果となった。これらのことから解説書による特徴と同じ結果を得ることができたと考えられる。一方で、ジグについても音価の大きい音の割合が高い結果となり、これは解説書による特徴と逆の結果が得ら

(2) n-gram を用いた構成音の分析結果

A phylogenetic tree with 100 taxa. The taxa are labeled as follows from top to bottom: set00001, set00002, set00003, set00004, set00005, set00006, set00007, set00008, set00009, set00010, set00011, set00012, set00013, set00014, set00015, set00016, set00017, set00018, set00019, set00020, set00021, set00022, set00023, set00024, set00025, set00026, set00027, set00028, set00029, set00030, set00031, set00032, set00033, set00034, set00035, set00036, set00037, set00038, set00039, set00040, set00041, set00042, set00043, set00044, set00045, set00046, set00047, set00048, set00049, set00050, set00051, set00052, set00053, set00054, set00055, set00056, set00057, set00058, set00059, set00060, set00061, set00062, set00063, set00064, set00065, set00066, set00067, set00068, set00069, set00070, set00071, set00072, set00073, set00074, set00075, set00076, set00077, set00078, set00079, set00080, set00081, set00082, set00083, set00084, set00085, set00086, set00087, set00088, set00089, set00090, set00091, set00092, set00093, set00094, set00095, set00096, set00097, set00098, set00099, set00100. The tree is rooted at the top left and branches downwards and to the right. The branching pattern is highly nested, with many taxa sharing common ancestors. The labels are in a monospace font.

[illegible]

音階の n-gram でのクラスタリングでは、 $n=2$ の時では同じ名前の楽曲や組曲でまとまることはなく、 n の値を 7 まで大きくすると同じ名前の楽曲が 3 曲まとまるなどの小さなまとまりがいくつか見られるようになった。一方、音価によるクラスタリングでは、 $n=2$ の時と $n=5$ の時で同じように大きく 2 つのまとまりに分かれる結果となり、楽曲の速さが速い曲と緩やかな曲に大雑把に分かれたものと考えられる。

(1) 式を用いてフレーズの分析を行なった結果の例として、組曲第 1 番のフレーズの分析結果を表 4 に、各前奏曲のフレーズに分析結果を表 5 に示す。

前奏曲	A, A', A', A', A', A', A', A', A', A', B
アルマンド	A, A', A', A', A', A', A', A', B
クーラント	A, A', A', A', A', A', A', A', A', A', A'
サラバンド	A, A', A', A'
メヌエット	A, A', A', A', A', A', B, B', B', B', B', B'
ジーク	A, A', A', A', A', A', A', B, B'

[illegible]

6. 分析結果の蓄積

分析を行なった音符やフレーズに関して、RDF (Resource Description Framework) の考えに基づいて、属性となる情報を3項組のデータとして蓄積を行う。RDFは、トリプルと呼ばれる構造によって関係を示すものであり、トリプルは事物の関係を主語、述語、目的語の関係を用いた表現を行う。RDF トリプルによる表現の例を図9に示す。

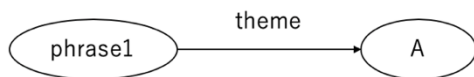


図 9 RDF トリプルによる表現例

図 9 では、「phrase1」が主語、「theme」が述語、「A」が目的語となり、「phrase1 は A という theme である」を示すものとなる。

RDF トリプルを利用し、以下の 3 つのデータの蓄積を行う。

- ・音階、音価の情報
- ・所属するフレーズと其中での位置
- ・音楽学での主題とその展開

結果として、組曲第 1 番前奏曲でのトリプルデータを表 6 に示す。

表 6 組曲第 1 番前奏曲でのトリプルデータ例

phrase1	theme	A
note1	step	A
note1	duration	0.25
note1	position	phrase1-1
note2	step	E

表 6 より、フレーズ 1 は「A」のテーマである。また、音符 1 は「A」の音で音価が 0.25 であり、フレーズ 1 の 1 番目の音であることを示している。

7. 考察

今回、分析結果から、構造化データである楽譜情報からの数理的処理による分析からでも解説書による特徴と同じ結果を得ることができると考えられる。また、フレーズ分析に使用した変数として、様々な楽曲で試行して算出した数値のようなより適切な値を使用することで、より客観的な分析結果を得ることができると考えられる。さらに、それぞれの楽曲において分析結果を得ることができたことから、ほかの楽曲に対して本手法を適用することは有効であり、知識グラフを充実させることでメロディー生成などほかの手法に応用させることができると考えられる。

8. 結論

本研究では、楽譜の構造化データである MusicXML を利用し、n-gram などの数理的処理を用いて音名や音価、フレーズについての分析を行うことができた。また、分析結果を RDF トリプルを利用した知識グラフとして蓄積することができた。

謝辞

本研究を進めていくにあたり、藤井章博教授には多くのご指導をいただき深く御礼申し上げます。また、ともに切磋琢磨してきた藤井研究室の皆様にも感謝いたします。

参考文献

- 1) 松崎裕佑, 梅村祥之: コード進行における非和声音に着目した主旋律の生成法開発および多様性の主観評価, 情報処理学会音楽情報科学 研究報告, Vol.2019-MUS-22, No1, 2019
- 2) 佐々木忠編: ギターのための無伴奏チェロ組曲, 全音楽譜出版社, 2001
- 3) 音楽之友社編: 最新名曲解説全集 第 14 巻 独奏曲 I, 音楽之友社, 1980
- 4) MusicXML 3.1 Tutorial, <http://wpmedia.musicxml.com/wp-content/uploads/2017/02/musicxml-tutorial.pdf>
- 5) 小谷野謙一: よくわかる楽典の教科書, 株式会社ヤマハミュージックメディア, 2011